

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (currently amended) A method of processing digitized textual information in a computerized database system, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, the method comprising:

generating a concept vector for each document in a document corpus wherein the concept vector conceptually classifying the contents of the document on a relatively compact format,

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors wherein the term-to-concept vectors being generated on basis of the concept vectors, comprises:

receiving the term-to-concept vectors for the document corpus and on basis thereof generating a term-term matrix describing a term-to-term relationship between the terms in the document corpus, and

processing the term-term matrix into processed textual information and displaying the processed textual information via a user output interface.

2. (previously presented) A method according to claim 1, wherein each document in the document corpus being associated with a document-concept matrix representing at least

one concept element whose relevance with respect to the document is described by a weight factor, the generation of each term-to-concept vector comprises:

- identifying a term-relevant set of documents in the document corpus, each document in the term-relevant set containing at least one occurrence of the term,
- calculating a term weight for the term in each of the documents in the term-relevant set,
- retrieving a respective concept vector being associated with each document in the term-relevant set where the term weight exceeds a first threshold value,
- selecting a relevant set of concept vectors including any concept vector in which at least one concept component exceeds a second threshold value,
- calculating a non-normalized term-to-concept vector as the sum of all concept vectors in the relevant set, and
- normalizing the non-normalized term-to-concept vector.

3. (previously presented) A method according to claim 1 wherein the generation of the term-term matrix comprises:

- retrieving, for each term in each combination of two unique terms in the document corpus, a respective term-to-concept vector,
- generating a relation vector describing the relationship between the terms in each combination of two unique terms, each component in the relation vector being equal to a lowest component value of corresponding component values in the term-to-concept vectors,

generating a relationship value for each combination of two unique terms as the sum of all component values in the corresponding relation vector, and

generating a matrix containing the relationship values of all combinations of two unique terms in the document corpus.

4. (previously presented) A method according to claim 1 wherein the method further comprises the steps of:

calculating a statistical co-occurrence value between each combination of two unique terms in the document corpus, the statistical co-occurrence value describing a dependent probability that a certain second term exists in a document provided that a certain first term exists in the document, and

incorporating the statistical co-occurrence values into the term-term matrix to represent lexical relationships between the terms in the document corpus.

5. (previously presented) A method according to claim 1 wherein the method further comprises the step of:

displaying the processed textual information on a format being adapted for human comprehension.

6. (previously presented) A method according to claim 5, wherein the displaying step further comprises involving presentation of:

at least one document identifier specifying a document being relevant with respect
at least one term in a query,

at least one term being related to a term in a query, and

a conceptual distribution representing a conceptual relationship between two or
more terms in the document corpus, the conceptual distribution being based on shared
concepts which are common to said terms.

7. (previously presented) A method according to claim 5 wherein the displaying step
further comprises involving presentation of at least one document identifier specifying a
document being relevant with respect to at least one term in a query in combination with
at least one user specified concept.

8. (previously presented) A method according to claim 6 wherein the method further
comprises the step of:

selecting the at least one user specified concept from the shared concepts in the
conceptual distribution.

9. (previously presented) A method according to claim 5 wherein the method further
comprises the step of:

illustrating the conceptual relationship between a first term and at least one
second term by means of a respective relevance measure being associated with the at
least one second term in respect of the first term.

10. (previously presented) A method according to claim 9, wherein the method further comprises the step of:

displaying the processed textual information on a graphical format which visualizes the strength in the conceptual relationship between at least two terms.

11. (previously presented) A method according to claim 9 wherein the method further comprises the steps of:

displaying the processed textual information as a distance graph in which each term constitutes a node wherein the anode representing a first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship of at least a specific strength, and the relevance measure between the first term and the at least one second term is represented by a minimum number of node hops between the first term and the at least one second term.

12. (previously presented) A method according to claim 9 wherein the method further comprises the step of:

displaying the processed textual information as a distance graph in which each term constitutes a node wherein the node representing a first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship, each connection is associated with an edge weight representing the strength of a conceptual relationship between the first term and a particular secondary term, and

the relevance measure between the first term and a particular secondary term is represented by an accumulation of the edge weights being associated with the connections constituting a minimum number node hops between the first term and the particular secondary term.

13. (previously presented) A method according to claim 1, wherein each term further comprises:

- a single word,
- a proper name,
- a phrase, and
- a compound of single words.

14. (previously presented) A method according to claim 1 further comprises the step of updating the document corpus with added data in form of at least one new document by means of

identifying any added terms in the new document which lack a representation in the document corpus,

identifying any existing terms in the new document which were represented in the document corpus before adding the at least one new document,

retrieving, for each of the existing terms, a corresponding concept vector,

generating a new concept vector with respect to the at least one new document as a sum of the corresponding concept vectors,

normalizing the new concept vector into a normalized new concept vector, and
assigning the normalized new concept vector to each of the added terms in the
new document.

15.-16. (canceled)

17. (currently amended) A computer-implemented search engine for processing an
amount of digitized textual information and extracting data there from, the information
being organized in terms, documents and document corpora, where each document
contains at least one term and each document corpus contains at least one document,
comprising:

an interface ~~adapted~~ configured to receive a query from a user, and
a processing unit ~~adapted~~ configured to process a document corpus on basis of the
query and return processed textual information being relevant to the query said process
involving

generating a concept vector for each document in the document corpus, the
concept vector conceptually classifying the contents of the document on a relatively'
compact format, and

generating, for each term in the document corpus, a term-to-concept vector
describing a relationship between the term and each of the concept vectors, wherein the
processing unit in turn comprises:

a processing module ~~adapted~~ configured to receive the term-to-concept vectors for the document corpus and on basis thereof generate a term-term matrix describing a term-to-term relation-ship between the terms in the document corpus, and

an exploring module ~~adapted~~ configured to receive the query and the term-term matrix, and on basis of the query process the term-term matrix into the processed textual information.

18. (currently amended) A computer-implemented database system holding an amount of digitized textual information being organized in terms, documents and document corpora, wherein each document contains at least one term and each document corpus contains at least one document, wherein each document in a document corpus being associated with concept vector which conceptually classifies the contents of the document on a relatively compact format, and wherein each term in the document corpus being associated with a term-to-concept vector describing a relationship between the term and each of the concept vectors, wherein it is ~~adapted~~ configured to deliver the term-to concept vectors to a search engine ~~according to claim 17. for processing an amount of digitized textual information and extracting data there from, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, comprising:~~

an interface configured to receive a query from a user, and

a processing unit configured to process a document corpus on basis of the query and return processed textual information being relevant to the query said process involving

generating a concept vector for each document in the document corpus, the concept vector conceptually classifying the contents of the document on a relatively' compact format, and

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors, wherein the processing unit in turn comprises:

a processing module configured to receive the term-to-concept vectors for the document corpus and on basis thereof generate a term-term matrix describing a term-to-term relation-ship between the terms in the document corpus, and

an exploring module configured to receive the query and the term-term matrix, and on basis of the query process the term-term matrix into the processed textual information.

19. (currently amended) A database system according to claim 18 further comprising an iterative term-to-concept engine ~~adapted~~ configured to receive fresh digitized textual information added to the database and on basis of this information

generate concept vectors for any added document, and

generate a term-to-concept vector describing a relationship between any added term and each of the concept vectors.

20. (currently amended) A server computer system for providing data processing services in respect of digitized textual information, wherein the server ~~comprises~~ comprising:

a search engine for processing an amount of digitized textual information and extracting data there from, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, comprising an interface ~~adapted~~ configured to receive a query from a user, and a processing unit ~~adapted~~ configured to process a document corpus on basis of the query and return processed textual information being relevant to the query said process involving generating a concept vector for each document in the document corpus, the concept vector conceptually classifying the contents of the document on a relatively compact format, and generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors, wherein the processing unit in turn comprises a processing module ~~adapted~~ configured to receive the term-to-concept vectors for the document corpus and on basis thereof generate a term-term matrix describing a term-to-term relation-ship between the terms in the document corpus, and an exploring module ~~adapted~~ configured to receive the query and the term-term matrix, and on basis of the query process the term-term matrix into the processed textual information and

a communication interface towards a database system according to claim 18.

21. (withdrawn) A system for providing data processing services in respect of digitized textual information, wherein the system comprising:

a server according to claim 20,

at least one user client adapted to communicate with the server, and

a communication link connecting the at least one user client with the server.

22. (withdrawn) A system according to claim 21 further comprising,

an internet accomplishes at least a part of the communication link, and the at least one user client comprises a web browser,

a user input interface adapted to receive queries from a user and forward the queries to the server via the communication link, and

a user output interface adapted to receive processed textual information from the server via the communication link and present the processed textual information to the user.

23. (withdrawn) A method of processing digitized textual information, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, the method comprising:

identifying a particular document corpus,

filtering the identified document corpus wherein a number of documents fulfilling at least one specified criterion are selected, and

producing a new document corpus exclusively containing the selected documents.

24. (withdrawn) A method according to claim 23 wherein, the filtering comprises the further steps of:

identifying a number of document clusters in the identified document corpus by means of a document clustering algorithm,

generating, for each identified document cluster, a representative document vector by means of the document clustering algorithm, and

removing all non-clustered documents from the identified document corpus.

25. (withdrawn) A method according to claim 23 wherein, the filtering comprises the further steps of:

receiving a user input specifying at least one of one or more concepts and one or more terms,

selecting, from the identified document corpus, documents being related to at least one of the concepts or the terms, and

removing all non-selected documents from the identified document corpus.

26. (withdrawn) A method according to claim 23, wherein the identified document corpus having been processed according to a method comprising the following steps:

generating a concept vector for each document in a document corpus wherein the concept vector conceptually classifying the contents of the document on a relatively compact format,

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors wherein the term-to-concept vectors being generated on basis of the concept vectors,

receiving the term-to-concept vectors for the document corpus and on basis thereof generating a term-term matrix describing a term-to-term relationship between the terms in the document corpus, and

processing the term-term matrix into processed textual information.

27. (new) A computer system comprising a processor for executing computer program instructions, a memory for storing computer program instructions and computer program instructions comprising software for performing a method of processing digitized textual information, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, the method comprising:

generating a concept vector for each document in a document corpus wherein the concept vector conceptually classifying the contents of the document on a relatively compact format,

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors wherein the term-to-concept vectors being generated on basis of the concept vectors,

receiving the term-to-concept vectors for the document corpus and on basis thereof generating a term-term matrix describing a term-to-term relationship between the terms in the document corpus, and

processing the term-term matrix into processed textual information and displaying the processed textual information via a user output interface.

28. (new) A computer program product comprising a computer readable storage medium, having computer program instructions recorded thereon for causing a computer to perform a method of processing digitized textual information, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, the method comprising:

generating a concept vector for each document in a document corpus wherein the concept vector conceptually classifying the contents of the document on a relatively compact format,

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors wherein the term-to-concept vectors being generated on basis of the concept vectors,

receiving the term-to-concept vectors for the document corpus and on basis thereof generating a term-term matrix describing a term-to-term relationship between the terms in the document corpus, and

processing the term-term matrix into processed textual information and displaying the processed textual information via a user output interface.